stats

MDPI

Article

General Formulas for the Central and Non-Central Moments of the Multinomial Distribution

Frédéric Ouimet

Department of Mathematics, California Institute of Technology, Pasadena, CA 91125, USA; ouimetfr@caltech.edu

Abstract: We present the first general formulas for the central and non-central moments of the multinomial distribution, using a combinatorial argument and the factorial moments previously obtained in Mosimann (1962). We use the formulas to give explicit expressions for all the non-central moments up to order 8 and all the central moments up to order 4. These results expand significantly on those in Newcomer (2008) and Newcomer et al. (2008), where the non-central moments were calculated up to order 4.

Keywords: multinomial distribution; higher moments; central moments; non-central moments

MSC: 62E15; 60E05

Citation: Ouimet, F. General Formulas for the Central and Non-Central Moments of the Multinomial Distribution. *Stats* **2021**, *4*, 18–27. https://doi.org/10.3390/stats4010002

Received: 10 November 2020
Accepted: 5 January 2021
Published: 6 January 2021

1. The Multinomial Distribution

For any $d \in \mathbb{N}$, the d-dimensional (unit) simplex is defined by $\mathcal{S} := \{\boldsymbol{x} \in [0,1]^d : \sum_{i=1}^d x_i \leq 1\}$, and the probability mass function $\boldsymbol{k} \mapsto P_{\boldsymbol{k},m}(\boldsymbol{x})$ for $\boldsymbol{\xi} := (\xi_1, \xi_2, \dots, \xi_d) \sim \mathrm{Multinomial}(m, \boldsymbol{x})$ is defined by

$$P_{\boldsymbol{k},m}(\boldsymbol{x}) := \frac{m!}{(m - \sum_{i=1}^d k_i)! \prod_{i=1}^d k_i!} \cdot (1 - \sum_{i=1}^d x_i)^{m - \sum_{i=1}^d k_i} \prod_{i=1}^d x_i^{k_i}, \quad \boldsymbol{k} \in \mathbb{N}_0^d \cap m\mathcal{S}, \tag{1}$$

where $m \in \mathbb{N}$ and $\boldsymbol{x} \in \mathcal{S}$. If $x_{d+1} := 1 - \sum_{i=1}^d x_i$, then (1) is just a reparametrization of $(\boldsymbol{\xi}, 1 - \sum_{i=1}^d \xi_i) \sim \mathrm{Multinomial}(m, (\boldsymbol{x}, x_{d+1}))$ where $\sum_{i=1}^{d+1} x_i = 1$. In this paper, our main goal is to give general formulas for the non-central and central moments of (1), namely

$$\mathbb{E}\Big[\prod_{i=1}^d \xi_i^{p_i}\Big] \quad \text{and} \quad \mathbb{E}\Big[\prod_{i=1}^d (\xi_i - \mathbb{E}[\xi_i])^{p_i}\Big], \qquad p_1, p_2, \dots, p_d \in \mathbb{N}_0. \tag{2}$$

We obtain the formulas using a combinatorial argument and the general expression for the factorial moments found in Mosimann (1962) [1], which we register in the lemma below.

Lemma 1 (Factorial moments). *Let* $\boldsymbol{\xi} \sim \mathrm{Multinomial}(m, \boldsymbol{x})$. *Then, for all* $p_1, p_2, \dots, p_d \in \mathbb{N}_0$,

$$\mathbb{E}\Big[\prod_{i=1}^d \xi_i^{(k_i)}\Big] = m^{(\sum_{i=1}^d k_i)} \prod_{i=1}^d x^{k_i}, \tag{3}$$

where $m^{(k)} := m(m-1)\dots(m-k+1)$ *denotes the k-th order falling factorial of m.*

The formulas that we develop for the expectations in (2) will be used to compute explicitly all the non-central moments up to order 8 and all the central moments up to order 4, which expands on the third and fourth order non-central moments that were previously calculated in (Newcomer (2008) [2], Appendix A.1). We should also mention that explicit formulas for several lower-order (mixed) cumulants were presented in Wishart (1949) [3] (see also Johnson et al. (1997) [4], page 37), but not for the moments.

2. Motivation

To the best of our knowledge, general formulas for the central and non-central moments of the multinomial distribution have never been derived in the literature. The central moments can arise naturally, for example, when studying asymptotic properties, via Taylor expansions of statistical estimators involving the multinomial distribution. For a given sequence of i.i.d. observations $\boldsymbol{X}_1, \boldsymbol{X}_2, \dots, \boldsymbol{X}_n$, two examples of such estimators are the Bernstein estimator for the cumulative distribution function

$$F_{n,m}^{\star}(\boldsymbol{x}) := \sum_{\boldsymbol{k} \in \mathbb{N}_0^d \cap m\mathcal{S}} \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{(-\boldsymbol{\infty}, \frac{\boldsymbol{k}}{m}]}(\boldsymbol{X}_i) P_{\boldsymbol{k},m}(\boldsymbol{x}), \quad \boldsymbol{x} \in \mathcal{S}, \ \ m, n \in \mathbb{N}, \tag{4}$$

and the Bernstein estimator for the density function (also called smoothed histogram)

$$\hat{f}_{n,m}(\boldsymbol{x}) := \sum_{\boldsymbol{k} \in \mathbb{N}_0^d \cap (m-1)\mathcal{S}} \frac{m^d}{n} \sum_{i=1}^{n} \mathbb{1}_{(\frac{\boldsymbol{k}}{m}, \frac{\boldsymbol{k}+1}{m}]}(\boldsymbol{X}_i) P_{\boldsymbol{k},m-1}(\boldsymbol{x}), \quad \boldsymbol{x} \in \mathcal{S}, \ \ m, n \in \mathbb{N}, \tag{5}$$

over the d-dimensional simplex. Some of their asymptotic properties were investigated by Stadtmüller (1986), Tenbusch (1997), Petrone (1999), Ghosal (2001), Petrone and Wasserman (2002), Babu et al. (2002), Kakizawa (2004), Bouezmarni and Rolin (2007), Bouezmarni et al. (2007), Leblanc (2009, 2010, 2012), Curtis and Ghosh (2011), Igarashi and Kakizawa(2014), Turnbull and Ghosh (2014), Lu (2015), Guan (2016, 2017) and Belalia et al. (2017, 2019) [5–30] when $d = 1$, by Tenbusch (1994) [31] when $d = 2$, and by Ouimet (2020) [32,33] for all $d \geq 1$, using a local limit theorem from Ouimet (2020) [34] for the multinomial distribution (see also Arenbaev (1976) [35]). The estimator (5) is a discrete analogue of the Dirichlet kernel estimator introduced by Aitchison and Lauder (1985) [36] and studied theoretically in Brown and Chen (1999), Chen (1999, 2000) and Bouezmarni and Rolin (2003 [37–40] when $d = 1$ (among others), and in Ouimet (2020) [41] for all $d \geq 1$.

3. Results

First, we give a general formula of the non-central moments of the multinomial distribution in (1).

Theorem 1 (Non-central moments)**.** *Let* $\boldsymbol{\xi} \sim \mathrm{Multinomial}(m, \boldsymbol{x})$*. For all* $p_1, p_2, \dots, p_d \in \mathbb{N}_0$*,*

$$\mathbb{E}\Big[\prod_{i=1}^{d} \xi_i^{p_i}\Big] = \sum_{k_1=0}^{p_1} \cdots \sum_{k_d=0}^{p_d} m^{(\sum_{i=1}^d k_i)} \prod_{i=1}^{d} \left\{ {p_i \atop k_i} \right\} x_i^{k_i}, \tag{6}$$

where $\left\{ {p \atop k} \right\}$ *denotes a Stirling number of the second kind (i.e., the number of ways to partition a set of p objects into k non-empty subsets).*

Proof. We have the following well-known relation between the power $p \in \mathbb{N}_0$ of a number $x \in \mathbb{R}$ and the falling factorials of x:

$$x^p = \sum_{k=0}^{p} \left\{ {p \atop k} \right\} x^{(k)}. \tag{7}$$

See, e.g., (Graham et al. (1994) [42], page 262). Apply this relation to every $\xi_i^{p_i}$ and use the linearity of the expectation to get

$$\mathbb{E}\Big[\prod_{i=1}^{d} \xi_i^{p_i}\Big] = \sum_{k_1=0}^{p_1} \cdots \sum_{k_d=0}^{p_d} \left\{ {p_1 \atop k_1} \right\} \cdots \left\{ {p_d \atop k_d} \right\} \mathbb{E}\Big[\prod_{i=1}^{d} \xi_i^{(k_i)}\Big]. \tag{8}$$

The conclusion follows from Lemma 1. □

We deduce a general formula for the central moments of the multinomial distribution.

Theorem 2 (Central moments)**.** *Let $\boldsymbol{\xi} \sim \mathrm{Multinomial}\,(m, \boldsymbol{x})$. For all $p_1, p_2, \ldots, p_d \in \mathbb{N}_0$,*

$$\mathbb{E}\Big[\prod_{i=1}^{d} (\xi_i - \mathbb{E}[\xi_i])^{p_i}\Big] = \sum_{\ell_1=0}^{p_1} \cdots \sum_{\ell_d=0}^{p_d} \sum_{k_1=0}^{\ell_1} \cdots \sum_{k_d=0}^{\ell_d} m^{(\sum_{i=1}^{d} k_i)} (-m)^{\sum_{i=1}^{d} (p_i - \ell_i)} \prod_{i=1}^{d} \binom{p_i}{\ell_i} {\ell_i \brace k_i} x_i^{p_i - \ell_i + k_i}, \tag{9}$$

where $\binom{p}{\ell}$ denotes the binomial coefficient $\frac{p!}{\ell!(p-\ell)!}$.

Proof. By applying the binomial formula to each factor $(\xi_i - \mathbb{E}[\xi_i])^{p_i}$ and using the fact that $\mathbb{E}[\xi_i] = m x_i$ for all $i \in \{1, 2, \ldots, d\}$, note that

$$\mathbb{E}\Big[\prod_{i=1}^{d} (\xi_i - \mathbb{E}[\xi_i])^{p_i}\Big] = \sum_{\ell_1=0}^{p_1} \cdots \sum_{\ell_d=0}^{p_d} \mathbb{E}\Big[\prod_{i=1}^{d} \xi_i^{\ell_i}\Big] \cdot \prod_{i=1}^{d} \binom{p_i}{\ell_i} (-m x_i)^{p_i - \ell_i}. \tag{10}$$

The conclusion follows from Theorem 1. □

4. Numerical Implementation

The formulas in Theorems 1 and 2 can be implemented in `Mathematica` as follows:

```
NonCentral[m_, x_, p_, d_] :=
  Sum[FactorialPower[m, Sum[k[i], {i,1,d}]] *
  Product[StirlingS2[p[[i]], k[i]] * x[[i]] ^ k[i], {i,1,d}], ##] & @@
  ({k[#], 0, p[[#]]} & /@ Range[d]);
Central[m_, x_, p_, d_] :=
  Sum[Sum[FactorialPower[m, Sum[k[i], {i,1,d}]] *
  (-m) ^ Sum[p[[i]] - ell[i], {i,1,d}] *
  Product[Binomial[p[[i]], ell[i]] * StirlingS2[ell[i], k[i]] *
  x[[i]] ^ (p[[i]] - ell[i] + k[i]), {i,1,d}], ##] & @@
  ({k[#], 0, ell[#]} & /@ Range[d]), ##] & @@
  ({ell[#], 0, p[[#]]} & /@ Range[d]);
```

5. Explicit Formulas

In Newcomer (2008) [2], explicit expressions for the non-central moments of order 3 and 4 were obtained for the multinomial distribution, see also Newcomer, Neerchal, and Morel (2008) and Ouimet (2020) [43,44]. To expand on those results, we use the formula from Theorem 1 in the two subsections below to calculate (explicitly) all the non-central moments up to order 8 and all the central moments up to order 4.

Here is a table of the Stirling numbers of the second kind that we will use in our calculations:

$$
\begin{aligned}
&\left\{{0 \atop 0}\right\} = 1,\\
&\left\{{1 \atop 0}\right\} = 0,\ \left\{{1 \atop 1}\right\} = 1,\\
&\left\{{2 \atop 0}\right\} = 0,\ \left\{{2 \atop 1}\right\} = 1, \left\{{2 \atop 2}\right\} = 1,\\
&\left\{{3 \atop 0}\right\} = 0,\ \left\{{3 \atop 1}\right\} = 1, \left\{{3 \atop 2}\right\} = 3,\ \left\{{3 \atop 3}\right\} = 1,\\
&\left\{{4 \atop 0}\right\} = 0,\ \left\{{4 \atop 1}\right\} = 1, \left\{{4 \atop 2}\right\} = 7,\ \left\{{4 \atop 3}\right\} = 6, \left\{{4 \atop 4}\right\} = 1,\\
&\left\{{5 \atop 0}\right\} = 0,\ \left\{{5 \atop 1}\right\} = 1, \left\{{5 \atop 2}\right\} = 15,\ \left\{{5 \atop 3}\right\} = 25, \left\{{5 \atop 4}\right\} = 10,\ \left\{{5 \atop 5}\right\} = 1,\\
&\left\{{6 \atop 0}\right\} = 0,\ \left\{{6 \atop 1}\right\} = 1, \left\{{6 \atop 2}\right\} = 31,\ \left\{{6 \atop 3}\right\} = 90, \left\{{6 \atop 4}\right\} = 65,\ \left\{{6 \atop 5}\right\} = 15, \left\{{6 \atop 6}\right\} = 1,\\
&\left\{{7 \atop 0}\right\} = 0,\ \left\{{7 \atop 1}\right\} = 1, \left\{{7 \atop 2}\right\} = 63,\ \left\{{7 \atop 3}\right\} = 301, \left\{{7 \atop 4}\right\} = 350,\ \left\{{7 \atop 5}\right\} = 140, \left\{{7 \atop 6}\right\} = 21,\ \left\{{7 \atop 7}\right\} = 1,\\
&\left\{{8 \atop 0}\right\} = 0,\ \left\{{8 \atop 1}\right\} = 1, \left\{{8 \atop 2}\right\} = 127,\ \left\{{8 \atop 3}\right\} = 966, \left\{{8 \atop 4}\right\} = 1701,\ \left\{{8 \atop 5}\right\} = 1050, \left\{{8 \atop 6}\right\} = 266,\ \left\{{8 \atop 7}\right\} = 28,\ \left\{{8 \atop 8}\right\} = 1.
\end{aligned}
$$

5.1. Computation of the Non-Central Moments up to Order 8

By applying the general expression in Theorem 1 and by removing the Stirling numbers $\{{p_i \atop k_i}\}$ that are equal to 0, we get the following results directly.

<u>Order 1:</u> For any $j_1 \in \{1, 2, \dots, d\}$,

$$\mathbb{E}[\xi_{j_1}] = x_{j_1} m. \tag{11}$$

<u>Order 2:</u> For any distinct $j_1, j_2 \in \{1, 2, \dots, d\}$,

$$\mathbb{E}[\xi_{j_1}^2] = x_{j_1} \left[m + m^{(2)} x_{j_1}\right], \tag{12}$$

$$\mathbb{E}[\xi_{j_1}\xi_{j_2}] = x_{j_1} x_{j_2} m^{(2)}. \tag{13}$$

<u>Order 3:</u> For any distinct $j_1, j_2, j_3 \in \{1, 2, \dots, d\}$,

$$\mathbb{E}[\xi_{j_1}^3] = x_{j_1} \left[m + 3m^{(2)} x_{j_1} + m^{(3)} x_{j_1}^2\right], \tag{14}$$

$$\mathbb{E}[\xi_{j_1}^2 \xi_{j_2}] = x_{j_1} x_{j_2} \left[m^{(2)} + m^{(3)} x_{j_1}\right], \tag{15}$$

$$\mathbb{E}[\xi_{j_1}\xi_{j_2}\xi_{j_3}] = x_{j_1} x_{j_2} x_{j_3} m^{(3)}. \tag{16}$$

<u>Order 4:</u> For any distinct $j_1, j_2, j_3, j_4 \in \{1, 2, \dots, d\}$,

$$\mathbb{E}[\xi_{j_1}^4] = x_{j_1} \left[m + 7m^{(2)} x_{j_1} + 6m^{(3)} x_{j_1}^2 + m^{(4)} x_{j_1}^3\right], \tag{17}$$

$$\mathbb{E}[\xi_{j_1}^3 \xi_{j_2}] = x_{j_1} x_{j_2} \left[m^{(2)} + 3m^{(3)} x_{j_1} + m^{(4)} x_{j_1}^2\right], \tag{18}$$

$$\mathbb{E}[\xi_{j_1}^2 \xi_{j_2}^2] = x_{j_1} x_{j_2} \left[m^{(2)} + m^{(3)} (x_{j_1} + x_{j_2}) + m^{(4)} x_{j_1} x_{j_2}\right], \tag{19}$$

$$\mathbb{E}[\xi_{j_1}^2 \xi_{j_2}\xi_{j_3}] = x_{j_1} x_{j_2} x_{j_3} \left[m^{(3)} + m^{(4)} x_{j_1}\right], \tag{20}$$

$$\mathbb{E}[\xi_{j_1}\xi_{j_2}\xi_{j_3}\xi_{j_4}] = x_{j_1} x_{j_2} x_{j_3} x_{j_4} m^{(4)}. \tag{21}$$

Order 5: For any distinct $j_1, j_2, j_3, j_4, j_5 \in \{1, 2, \ldots, d\}$,

$$\mathbb{E}[\xi_{j_1}^5] = x_{j_1}\left[m + 15m^{(2)}x_{j_1} + 25m^{(3)}x_{j_1}^2 + 10m^{(4)}x_{j_1}^3 + m^{(5)}x_{j_1}^4\right], \tag{22}$$

$$\mathbb{E}[\xi_{j_1}^4\xi_{j_2}] = x_{j_1}x_{j_2}\left[m^{(2)} + 7m^{(3)}x_{j_1} + 6m^{(4)}x_{j_1}^2 + m^{(5)}x_{j_1}^3\right], \tag{23}$$

$$\mathbb{E}[\xi_{j_1}^3\xi_{j_2}^2] = x_{j_1}x_{j_2}\left[m^{(2)} + m^{(3)}(3x_{j_1} + x_{j_2}) + m^{(4)}(x_{j_1}^2 + 3x_{j_1}x_{j_2}) + m^{(5)}x_{j_1}^2x_{j_2}\right], \tag{24}$$

$$\mathbb{E}[\xi_{j_1}^3\xi_{j_2}\xi_{j_3}] = x_{j_1}x_{j_2}x_{j_3}\left[m^{(3)} + 3m^{(4)}x_{j_1} + m^{(5)}x_{j_1}^2\right], \tag{25}$$

$$\mathbb{E}[\xi_{j_1}^2\xi_{j_2}^2\xi_{j_3}] = x_{j_1}x_{j_2}x_{j_3}\left[m^{(3)} + m^{(4)}(x_{j_1} + x_{j_2}) + m^{(5)}x_{j_1}x_{j_2}\right], \tag{26}$$

$$\mathbb{E}[\xi_{j_1}^2\xi_{j_2}\xi_{j_3}\xi_{j_4}] = x_{j_1}x_{j_2}x_{j_3}x_{j_4}\left[m^{(4)} + m^{(5)}x_{j_1}\right], \tag{27}$$

$$\mathbb{E}[\xi_{j_1}\xi_{j_2}\xi_{j_3}\xi_{j_4}\xi_{j_5}] = x_{j_1}x_{j_2}x_{j_3}x_{j_4}x_{j_5}m^{(5)}. \tag{28}$$

Order 6: For any distinct $j_1, j_2, j_3, j_4, j_5, j_6 \in \{1, 2, \ldots, d\}$,

$$\mathbb{E}[\xi_{j_1}^6] = x_{j_1}\left[m + 31m^{(2)}x_{j_1} + 90m^{(3)}x_{j_1}^2 + 65m^{(4)}x_{j_1}^3 + 15m^{(5)}x_{j_1}^4 + m^{(6)}x_{j_1}^5\right], \tag{29}$$

$$\mathbb{E}[\xi_{j_1}^5\xi_{j_2}] = x_{j_1}x_{j_2}\left[m^{(2)} + 15m^{(3)}x_{j_1} + 25m^{(4)}x_{j_1}^2 + 10m^{(5)}x_{j_1}^3 + m^{(6)}x_{j_1}^4\right], \tag{30}$$

$$\mathbb{E}[\xi_{j_1}^4\xi_{j_2}^2] = x_{j_1}x_{j_2}\begin{bmatrix} m^{(2)} + m^{(3)}(7x_{j_1} + x_{j_2}) + m^{(4)}(6x_{j_1}^2 + 7x_{j_1}x_{j_2}) \\ +m^{(5)}(x_{j_1}^3 + 6x_{j_1}^2x_{j_2}) + m^{(6)}x_{j_1}^3x_{j_2} \end{bmatrix}, \tag{31}$$

$$\mathbb{E}[\xi_{j_1}^4\xi_{j_2}\xi_{j_3}] = x_{j_1}x_{j_2}x_{j_3}\left[m^{(3)} + 7m^{(4)}x_{j_1} + 6m^{(5)}x_{j_1}^2 + m^{(6)}x_{j_1}^3\right], \tag{32}$$

$$\mathbb{E}[\xi_{j_1}^3\xi_{j_2}^3] = x_{j_1}x_{j_2}\begin{bmatrix} m^{(2)} + m^{(3)}(3x_{j_1} + 3x_{j_2}) + m^{(4)}(x_{j_1}^2 + 9x_{j_1}x_{j_2} + x_{j_2}^2) \\ +m^{(5)}(3x_{j_1}^2x_{j_2} + 3x_{j_1}x_{j_2}^2) + m^{(6)}x_{j_1}^2x_{j_2}^2 \end{bmatrix}, \tag{33}$$

$$\mathbb{E}[\xi_{j_1}^3\xi_{j_2}^2\xi_{j_3}] = x_{j_1}x_{j_2}x_{j_3}\left[m^{(3)} + m^{(4)}(3x_{j_1} + x_{j_2}) + m^{(5)}(x_{j_1}^2 + 3x_{j_1}x_{j_2}) + m^{(6)}x_{j_1}^2x_{j_2}\right], \tag{34}$$

$$\mathbb{E}[\xi_{j_1}^3\xi_{j_2}\xi_{j_3}\xi_{j_4}] = x_{j_1}x_{j_2}x_{j_3}x_{j_4}\left[m^{(4)} + 3m^{(5)}x_{j_1} + m^{(6)}x_{j_1}^2\right], \tag{35}$$

$$\mathbb{E}[\xi_{j_1}^2\xi_{j_2}^2\xi_{j_3}^2] = x_{j_1}x_{j_2}x_{j_3}\begin{bmatrix} m^{(3)} + m^{(4)}(x_{j_1} + x_{j_2} + x_{j_3}) \\ +m^{(5)}(x_{j_1}x_{j_2} + x_{j_1}x_{j_3} + x_{j_2}x_{j_3}) + m^{(6)}x_{j_1}x_{j_2}x_{j_3} \end{bmatrix}, \tag{36}$$

$$\mathbb{E}[\xi_{j_1}^2\xi_{j_2}^2\xi_{j_3}\xi_{j_4}] = x_{j_1}x_{j_2}x_{j_3}x_{j_4}\left[m^{(4)} + m^{(5)}(x_{j_1} + x_{j_2}) + m^{(6)}x_{j_1}x_{j_2}\right], \tag{37}$$

$$\mathbb{E}[\xi_{j_1}^2\xi_{j_2}\xi_{j_3}\xi_{j_4}\xi_{j_5}] = x_{j_1}x_{j_2}x_{j_3}x_{j_4}x_{j_5}\left[m^{(5)} + m^{(6)}x_{j_1}\right], \tag{38}$$

$$\mathbb{E}[\xi_{j_1}\xi_{j_2}\xi_{j_3}\xi_{j_4}\xi_{j_5}\xi_{j_6}] = x_{j_1}x_{j_2}x_{j_3}x_{j_4}x_{j_5}x_{j_6}m^{(6)}. \tag{39}$$

Order 7: For any distinct $j_1, j_2, j_3, j_4, j_5, j_6, j_7 \in \{1, 2, \ldots, d\}$,

$$\mathbb{E}[\xi_{j_1}^7] = x_{j_1}\begin{bmatrix} m + 63m^{(2)}x_{j_1} + 301m^{(3)}x_{j_1}^2 + 350m^{(4)}x_{j_1}^3 \\ +140m^{(5)}x_{j_1}^4 + 21m^{(6)}x_{j_1}^5 + m^{(7)}x_{j_1}^6 \end{bmatrix}, \tag{40}$$

$$\mathbb{E}[\xi_{j_1}^6\xi_{j_2}] = x_{j_1}x_{j_2}\begin{bmatrix} m^{(2)} + 31m^{(3)}x_{j_1} + 90m^{(4)}x_{j_1}^2 \\ +65m^{(5)}x_{j_1}^3 + 15m^{(6)}x_{j_1}^4 + m^{(7)}x_{j_1}^5 \end{bmatrix}, \tag{41}$$

$$\mathbb{E}[\xi_{j_1}^5\xi_{j_2}^2] = x_{j_1}x_{j_2}\begin{bmatrix} m^{(2)} + m^{(3)}(15x_{j_1} + x_{j_2}) + m^{(4)}(25x_{j_1}^2 + 15x_{j_1}x_{j_2}) \\ +m^{(5)}(10x_{j_1}^3 + 25x_{j_1}^2x_{j_2}) + m^{(6)}(x_{j_1}^4 + 10x_{j_1}^3x_{j_2}) + m^{(7)}x_{j_1}^4x_{j_2} \end{bmatrix}, \tag{42}$$

$$\mathbb{E}[\xi_{j_1}^5\xi_{j_2}\xi_{j_3}] = x_{j_1}x_{j_2}x_{j_3}\left[m^{(3)} + 15m^{(4)}x_{j_1} + 25m^{(5)}x_{j_1}^2 + 10m^{(6)}x_{j_1}^3 + m^{(7)}x_{j_1}^4\right], \tag{43}$$

$$\mathbb{E}[\xi_{j_1}^4\xi_{j_2}^3] = x_{j_1}x_{j_2}\begin{bmatrix} m^{(2)} + m^{(3)}(7x_{j_1} + 3x_{j_2}) + m^{(4)}(6x_{j_1}^2 + 21x_{j_1}x_{j_2} + x_{j_2}^2) \\ +m^{(5)}(x_{j_1}^3 + 18x_{j_1}^2x_{j_2} + 7x_{j_1}x_{j_2}^2) \\ +m^{(6)}(3x_{j_1}^3x_{j_2} + 6x_{j_1}^2x_{j_2}^2) + m^{(7)}x_{j_1}^3x_{j_2}^2 \end{bmatrix}, \tag{44}$$

$$\mathbb{E}[\xi_{j_1}^4\xi_{j_2}^2\xi_{j_3}] = x_{j_1}x_{j_2}x_{j_3}\begin{bmatrix} m^{(3)} + m^{(4)}(7x_{j_1} + x_{j_2}) + m^{(5)}(6x_{j_1}^2 + 7x_{j_1}x_{j_2}) \\ +m^{(6)}(x_{j_1}^3 + 6x_{j_1}^2x_{j_2}) + m^{(7)}x_{j_1}^3x_{j_2} \end{bmatrix}, \tag{45}$$

$$\mathbb{E}[\xi_{j_1}^4\xi_{j_2}\xi_{j_3}\xi_{j_4}] = x_{j_1}x_{j_2}x_{j_3}x_{j_4}[m^{(4)} + 7m^{(5)}x_{j_1} + 6m^{(6)}x_{j_1}^2 + m^{(7)}x_{j_1}^3], \tag{46}$$

$$\mathbb{E}[\xi_{j_1}^3\xi_{j_2}^3\xi_{j_3}] = x_{j_1}x_{j_2}x_{j_3}\begin{bmatrix} m^{(3)} + m^{(4)}(3x_{j_1} + 3x_{j_2}) + m^{(5)}(x_{j_1}^2 + 9x_{j_1}x_{j_2} + x_{j_2}^2) \\ +m^{(6)}(3x_{j_1}^2x_{j_2} + 3x_{j_1}x_{j_2}^2) + m^{(7)}x_{j_1}^2x_{j_2}^2 \end{bmatrix}, \tag{47}$$

$$\mathbb{E}[\xi_{j_1}^3\xi_{j_2}^2\xi_{j_3}^2] = x_{j_1}x_{j_2}x_{j_3}\begin{bmatrix} m^{(3)} + m^{(4)}(3x_{j_1} + x_{j_2} + x_{j_3}) \\ +m^{(5)}(x_{j_1}^2 + 3x_{j_1}x_{j_2} + 3x_{j_1}x_{j_3} + x_{j_2}x_{j_3}) \\ +m^{(6)}(x_{j_1}^2x_{j_2} + x_{j_1}^2x_{j_3} + 3x_{j_1}x_{j_2}x_{j_3}) + m^{(7)}x_{j_1}^2x_{j_2}x_{j_3} \end{bmatrix}, \tag{48}$$

$$\mathbb{E}[\xi_{j_1}^3\xi_{j_2}^2\xi_{j_3}\xi_{j_4}] = x_{j_1}x_{j_2}x_{j_3}x_{j_4}\begin{bmatrix} m^{(4)} + m^{(5)}(3x_{j_1} + x_{j_2}) \\ +m^{(6)}(x_{j_1}^2 + 3x_{j_1}x_{j_2}) + m^{(7)}x_{j_1}^2x_{j_2} \end{bmatrix}, \tag{49}$$

$$\mathbb{E}[\xi_{j_1}^3\xi_{j_2}\xi_{j_3}\xi_{j_4}\xi_{j_5}] = x_{j_1}x_{j_2}x_{j_3}x_{j_4}x_{j_5}[m^{(5)} + 3m^{(6)}x_{j_1} + m^{(7)}x_{j_1}^2], \tag{50}$$

$$\mathbb{E}[\xi_{j_1}^2\xi_{j_2}^2\xi_{j_3}^2\xi_{j_4}] = x_{j_1}x_{j_2}x_{j_3}x_{j_4}\begin{bmatrix} m^{(4)} + m^{(5)}(x_{j_1} + x_{j_2} + x_{j_3}) \\ +m^{(6)}(x_{j_1}x_{j_2} + x_{j_1}x_{j_3} + x_{j_2}x_{j_3}) + m^{(7)}x_{j_1}x_{j_2}x_{j_3} \end{bmatrix}, \tag{51}$$

$$\mathbb{E}[\xi_{j_1}^2\xi_{j_2}^2\xi_{j_3}\xi_{j_4}\xi_{j_5}] = x_{j_1}x_{j_2}x_{j_3}x_{j_4}x_{j_5}[m^{(5)} + m^{(6)}(x_{j_1} + x_{j_2}) + m^{(7)}x_{j_1}x_{j_2}], \tag{52}$$

$$\mathbb{E}[\xi_{j_1}^2\xi_{j_2}\xi_{j_3}\xi_{j_4}\xi_{j_5}\xi_{j_6}] = x_{j_1}x_{j_2}x_{j_3}x_{j_4}x_{j_5}x_{j_6}[m^{(6)} + m^{(7)}x_{j_1}], \tag{53}$$

$$\mathbb{E}[\xi_{j_1}\xi_{j_2}\xi_{j_3}\xi_{j_4}\xi_{j_5}\xi_{j_6}\xi_{j_7}] = x_{j_1}x_{j_2}x_{j_3}x_{j_4}x_{j_5}x_{j_6}x_{j_7}m^{(7)}. \tag{54}$$

<u>Order 8:</u> For any distinct $j_1, j_2, j_3, j_4, j_5, j_6, j_7, j_8 \in \{1, 2, \ldots, d\}$,

$$\mathbb{E}[\xi_{j_1}^8] = x_{j_1}\begin{bmatrix} m + 127m^{(2)}x_{j_1} + 966m^{(3)}x_{j_2}^2 + 1701m^{(4)}x_{j_1}^3 \\ +1050m^{(5)}x_{j_1}^4 + 266m^{(6)}x_{j_1}^5 + 28m^{(7)}x_{j_1}^6 + m^{(8)}x_{j_1}^7 \end{bmatrix}, \tag{55}$$

$$\mathbb{E}[\xi_{j_1}^7\xi_{j_2}] = x_{j_1}x_{j_2}\begin{bmatrix} m + 63m^{(3)}x_{j_1} + 301m^{(4)}x_{j_1}^2 + 350m^{(5)}x_{j_1}^3 \\ +140m^{(6)}x_{j_1}^4 + 21m^{(7)}x_{j_1}^5 + m^{(8)}x_{j_1}^6 \end{bmatrix}, \tag{56}$$

$$\mathbb{E}[\xi_{j_1}^6\xi_{j_2}^2] = x_{j_1}x_{j_2}\begin{bmatrix} m^{(2)} + m^{(3)}(31x_{j_1} + x_{j_2}) + m^{(4)}(90x_{j_1}^2 + 31x_{j_1}x_{j_2}) \\ +m^{(5)}(65x_{j_1}^3 + 90x_{j_1}^2x_{j_2}) + m^{(6)}(15x_{j_1}^4 + 65x_{j_1}^3x_{j_2}) \\ +m^{(7)}(x_{j_1}^5 + 15x_{j_1}^4x_{j_2}) + m^{(8)}x_{j_1}^5x_{j_2} \end{bmatrix}, \tag{57}$$

$$\mathbb{E}[\xi_{j_1}^6\xi_{j_2}\xi_{j_3}] = x_{j_1}x_{j_2}x_{j_3}\begin{bmatrix} m^{(3)} + 31m^{(4)}x_{j_1} + 90m^{(5)}x_{j_1}^2 \\ +65m^{(6)}x_{j_1}^3 + 15m^{(7)}x_{j_1}^4 + m^{(8)}x_{j_1}^5 \end{bmatrix}, \tag{58}$$

$$\mathbb{E}[\xi_{j_1}^5\xi_{j_2}^3] = x_{j_1}x_{j_2}\begin{bmatrix} m^{(2)} + m^{(3)}(15x_{j_1} + 3x_{j_2}) + m^{(4)}(25x_{j_1}^2 + 45x_{j_1}x_{j_2} + x_{j_2}^2) \\ +m^{(5)}(10x_{j_1}^3 + 75x_{j_1}^2x_{j_2} + 15x_{j_1}x_{j_2}^2) \\ +m^{(6)}(x_{j_4}^4 + 30x_{j_1}^3x_{j_2} + 25x_{j_1}^2x_{j_2}^2) \\ +m^{(7)}(3x_{j_1}^4x_{j_2} + 10x_{j_1}^3x_{j_2}^2) + m^{(8)}x_{j_1}^4x_{j_2}^2 \end{bmatrix}, \tag{59}$$

$$\mathbb{E}[\xi_{j_1}^5\xi_{j_2}^2\xi_{j_3}] = x_{j_1}x_{j_2}x_{j_3}\begin{bmatrix} m^{(3)} + m^{(4)}(15x_{j_1} + x_{j_2}) + m^{(5)}(25x_{j_1}^2 + 15x_{j_1}x_{j_2}) \\ +m^{(6)}(10x_{j_1}^3 + 25x_{j_1}^2x_{j_2}) + m^{(7)}(x_{j_1}^4 + 10x_{j_1}^3x_{j_2}) + m^{(8)}x_{j_1}^4x_{j_2} \end{bmatrix}, \tag{60}$$

$$\mathbb{E}[\xi_{j_1}^5\xi_{j_2}\xi_{j_3}\xi_{j_4}] = x_{j_1}x_{j_2}x_{j_3}x_{j_4}[m^{(4)} + 15m^{(5)}x_{j_1} + 25m^{(6)}x_{j_1}^2 + 10m^{(7)}x_{j_1}^3 + m^{(8)}x_{j_1}^4], \tag{61}$$

$$\mathbb{E}[\xi_{j_1}^4\xi_{j_2}^4] = x_{j_1}x_{j_2}\begin{bmatrix} m^{(2)} + m^{(3)}(7x_{j_1} + 7x_{j_2}) + m^{(4)}(6x_{j_1}^2 + 49x_{j_1}x_{j_2} + 6x_{j_2}^2) \\ +m^{(5)}(x_{j_1}^3 + 42x_{j_1}^2x_{j_2} + 42x_{j_1}x_{j_2}^2 + x_{j_2}^3) \\ +m^{(6)}(7x_{j_1}^3x_{j_2} + 36x_{j_1}^2x_{j_2}^2 + 7x_{j_1}x_{j_2}^3) \\ +m^{(7)}(6x_{j_1}^3x_{j_2}^2 + 6x_{j_1}^2x_{j_2}^3) + m^{(8)}x_{j_1}^3x_{j_2}^3 \end{bmatrix}, \tag{62}$$

$$\mathbb{E}[\xi_{j_1}^4\xi_{j_2}^3\xi_{j_3}] = x_{j_1}x_{j_2}x_{j_3}\begin{bmatrix} m^{(3)} + m^{(4)}(7x_{j_1} + 3x_{j_2}) + m^{(5)}(6x_{j_1}^2 + 21x_{j_1}x_{j_2} + x_{j_2}^2) \\ +m^{(6)}(x_{j_1}^3 + 18x_{j_1}^2x_{j_2} + 7x_{j_1}x_{j_2}^2) \\ +m^{(7)}(3x_{j_1}^3x_{j_2} + 6x_{j_1}^2x_{j_2}^2) + m^{(8)}x_{j_1}^3x_{j_2}^2 \end{bmatrix}, \tag{63}$$

$$\mathbb{E}[\xi_{j_1}^4\xi_{j_2}^2\xi_{j_3}^2] = x_{j_1}x_{j_2}x_{j_3}\begin{bmatrix} m^{(3)} + m^{(4)}(7x_{j_1} + x_{j_2} + x_{j_3}) \\ +m^{(5)}(6x_{j_1}^2 + 7x_{j_1}x_{j_2} + 7x_{j_1}x_{j_3} + x_{j_2}x_{j_3}) \\ +m^{(6)}(x_{j_1}^3 + 6x_{j_1}^2x_{j_2} + 6x_{j_1}^2x_{j_3} + 7x_{j_1}x_{j_2}x_{j_3}) \\ +m^{(7)}(x_{j_1}^3x_{j_2} + x_{j_1}^3x_{j_3} + 6x_{j_1}^2x_{j_2}x_{j_3}) + m^{(8)}x_{j_1}^3x_{j_2}x_{j_3} \end{bmatrix}, \tag{64}$$

$$\mathbb{E}[\xi_{j_1}^4\xi_{j_2}^2\xi_{j_3}\xi_{j_4}] = x_{j_1}x_{j_2}x_{j_3}x_{j_4}\begin{bmatrix} m^{(4)} + m^{(5)}(7x_{j_1} + x_{j_2}) + m^{(6)}(6x_{j_1}^2 + 7x_{j_1}x_{j_2}) \\ +m^{(7)}(x_{j_1}^3 + 6x_{j_1}^2x_{j_2}) + m^{(8)}x_{j_1}^3x_{j_2} \end{bmatrix}, \tag{65}$$

$$\mathbb{E}[\xi_{j_1}^4\xi_{j_2}\xi_{j_3}\xi_{j_4}\xi_{j_5}] = x_{j_1}x_{j_2}x_{j_3}x_{j_4}x_{j_5}\left[m^{(5)} + 7m^{(6)}x_{j_1} + 6m^{(7)}x_{j_1}^2 + m^{(8)}x_{j_1}^3\right], \tag{66}$$

$$\mathbb{E}[\xi_{j_1}^3\xi_{j_2}^3\xi_{j_3}^2] = x_{j_1}x_{j_2}x_{j_3}\begin{bmatrix} m^{(3)} + m^{(4)}(3x_{j_1} + 3x_{j_2} + x_{j_3}) \\ +m^{(5)}(x_{j_1}^2 + x_{j_2}^2 + 3x_{j_1}x_{j_3} + 3x_{j_2}x_{j_3} + 9x_{j_1}x_{j_2}) \\ +m^{(6)}(x_{j_1}^2x_{j_3} + x_{j_2}^2x_{j_3} + 3x_{j_1}^2x_{j_2} + 3x_{j_1}x_{j_2}^2 + 9x_{j_1}x_{j_2}x_{j_3}) \\ +m^{(7)}(x_{j_1}^2x_{j_2}^2 + 3x_{j_1}^2x_{j_2}x_{j_3} + 3x_{j_1}x_{j_2}^2x_{j_3}) + m^{(8)}x_{j_1}^2x_{j_2}^2x_{j_3} \end{bmatrix}, \tag{67}$$

$$\mathbb{E}[\xi_{j_1}^3\xi_{j_2}^3\xi_{j_3}\xi_{j_4}] = x_{j_1}x_{j_2}x_{j_3}x_{j_4}\begin{bmatrix} m^{(4)} + m^{(5)}(3x_{j_1} + 3x_{j_2}) + m^{(6)}(x_{j_1}^2 + 9x_{j_1}x_{j_2} + x_{j_2}^2) \\ +m^{(7)}(3x_{j_1}^2x_{j_2} + 3x_{j_1}x_{j_2}^2) + m^{(8)}x_{j_1}^2x_{j_2}^2 \end{bmatrix}, \tag{68}$$

$$\mathbb{E}[\xi_{j_1}^3\xi_{j_2}^2\xi_{j_3}^2\xi_{j_4}] = x_{j_1}x_{j_2}x_{j_3}x_{j_4}\begin{bmatrix} m^{(4)} + m^{(5)}(3x_{j_1} + x_{j_2} + x_{j_3}) \\ +m^{(6)}(3x_{j_1}x_{j_2} + 3x_{j_1}x_{j_3} + x_{j_2}x_{j_3}) \\ +m^{(7)}(x_{j_1}^2x_{j_2} + x_{j_1}^2x_{j_3} + 3x_{j_1}x_{j_2}x_{j_3}) + m^{(8)}x_{j_1}^2x_{j_2}x_{j_3} \end{bmatrix}, \tag{69}$$

$$\mathbb{E}[\xi_{j_1}^3\xi_{j_2}^2\xi_{j_3}\xi_{j_4}\xi_{j_5}] = x_{j_1}x_{j_2}x_{j_3}x_{j_4}x_{j_5}\begin{bmatrix} m^{(5)} + m^{(6)}(3x_{j_1} + x_{j_2}) \\ +m^{(7)}(x_{j_1}^2 + 3x_{j_1}x_{j_2}) + m^{(8)}x_{j_1}^2x_{j_2} \end{bmatrix}, \tag{70}$$

$$\mathbb{E}[\xi_{j_1}^3\xi_{j_2}\xi_{j_3}\xi_{j_4}\xi_{j_5}\xi_{j_6}] = x_{j_1}x_{j_2}x_{j_3}x_{j_4}x_{j_5}x_{j_6}\left[m^{(6)} + 3m^{(7)}x_{j_1} + m^{(8)}x_{j_1}^2\right], \tag{71}$$

$$\mathbb{E}[\xi_{j_1}^2\xi_{j_2}^2\xi_{j_3}^2\xi_{j_4}^2] = x_{j_1}x_{j_2}x_{j_3}x_{j_4}\begin{bmatrix} m^{(4)} + m^{(5)}(x_{j_1} + x_{j_2} + x_{j_3} + x_{j_4}) \\ +m^{(6)}(x_{j_1}x_{j_2} + x_{j_1}x_{j_3} + x_{j_1}x_{j_4} + x_{j_2}x_{j_3} + x_{j_2}x_{j_4} + x_{j_3}x_{j_4}) \\ +m^{(7)}(x_{j_1}x_{j_2}x_{j_3} + x_{j_1}x_{j_2}x_{j_4} + x_{j_1}x_{j_3}x_{j_4} + x_{j_2}x_{j_3}x_{j_4}) \\ +m^{(8)}x_{j_1}x_{j_2}x_{j_3}x_{j_4} \end{bmatrix}, \tag{72}$$

$$\mathbb{E}[\xi_{j_1}^2\xi_{j_2}^2\xi_{j_3}^2\xi_{j_4}\xi_{j_5}] = x_{j_1}x_{j_2}x_{j_3}x_{j_4}x_{j_5}\begin{bmatrix} m^{(5)} + m^{(6)}(x_{j_1} + x_{j_2} + x_{j_3}) \\ +m^{(7)}(x_{j_1}x_{j_2} + x_{j_1}x_{j_3} + x_{j_2}x_{j_3}) + m^{(8)}x_{j_1}x_{j_2}x_{j_3} \end{bmatrix}, \tag{73}$$

$$\mathbb{E}[\xi_{j_1}^2\xi_{j_2}^2\xi_{j_3}\xi_{j_4}\xi_{j_5}\xi_{j_6}] = x_{j_1}x_{j_2}x_{j_3}x_{j_4}x_{j_5}x_{j_6}\left[m^{(6)} + m^{(7)}(x_{j_1} + x_{j_2}) + m^{(8)}x_{j_1}x_{j_2}\right], \tag{74}$$

$$\mathbb{E}[\xi_{j_1}^2\xi_{j_2}\xi_{j_3}\xi_{j_4}\xi_{j_5}\xi_{j_6}\xi_{j_7}] = x_{j_1}x_{j_2}x_{j_3}x_{j_4}x_{j_5}x_{j_6}x_{j_7}\left[m^{(7)} + m^{(8)}x_{j_1}\right], \tag{75}$$

$$\mathbb{E}[\xi_{j_1}\xi_{j_2}\xi_{j_3}\xi_{j_4}\xi_{j_5}\xi_{j_6}\xi_{j_7}\xi_{j_8}] = x_{j_1}x_{j_2}x_{j_3}x_{j_4}x_{j_5}x_{j_6}x_{j_7}x_{j_8}m^{(8)}. \tag{76}$$

5.2. Computation of the Central Moments up to Order 4

With the results of the previous subsection and some algebraic manipulations (or the formula in Theorem 2), we can now calculate the central moments explicitly. We could calculate them up to order 8, but it would be very tedious. Instead, we write them up to order 4 for the sake of brevity. The simplifications we make to obtain the boxed expressions below are done with `Mathematica`.

Order 2: For any distinct $j_1, j_2 \in \{1, 2, \dots, d\}$,

$$\begin{aligned} \mathbb{E}[(\xi_{j_1} - \mathbb{E}[\xi_{j_1}])^2] &= \mathbb{E}[\xi_{j_1}^2] - (\mathbb{E}[\xi_{j_1}])^2 = x_{j_1}\left[m + m^{(2)}x_{j_1}\right] - m^2x_{j_1}^2 \\ &= \boxed{mx_{j_1}(1 - x_{j_1})} \end{aligned} \tag{77}$$

$$\begin{aligned} \mathbb{E}[(\xi_{j_1} - \mathbb{E}[\xi_{j_1}])(\xi_{j_2} - \mathbb{E}[\xi_{j_2}])] &= \mathbb{E}[\xi_{j_1}\xi_{j_2}] - \mathbb{E}[\xi_{j_1}]\mathbb{E}[\xi_{j_2}] = m^{(2)}x_{j_1}x_{j_2} - mx_{j_1}mx_{j_2} \\ &= \boxed{-mx_{j_1}x_{j_2}}. \end{aligned} \tag{78}$$

<u>Order 3:</u> For any distinct $j_1, j_2, j_3 \in \{1, 2, \dots, d\}$,

$$\begin{aligned}
\mathbb{E}[(\xi_{j_1} - \mathbb{E}[\xi_{j_1}])^3] &= \mathbb{E}[\xi_{j_1}^3] - 3\,\mathbb{E}[\xi_{j_1}^2]\mathbb{E}[\xi_{j_1}] + 2\,(\mathbb{E}[\xi_{j_1}])^3 \\
&= x_{j_1}\left[m + 3m^{(2)}x_{j_1} + m^{(3)}x_{j_1}^2\right] - 3x_{j_1}\left[m + m^{(2)}x_{j_1}\right]mx_{j_1} + 2m^3x_{j_1}^3 \\
&= \boxed{mx_{j_1}(x_{j_1} - 1)(2x_{j_1} - 1)} \qquad (79)
\end{aligned}$$

$$\begin{aligned}
&\mathbb{E}[(\xi_{j_1} - \mathbb{E}[\xi_{j_1}])^2(\xi_{j_2} - \mathbb{E}[\xi_{j_2}])] \\
&\quad = \mathbb{E}[\xi_{j_1}^2\xi_{j_2}] - \mathbb{E}[\xi_{j_1}^2]\mathbb{E}[\xi_{j_2}] - 2\,\mathbb{E}[\xi_{j_1}\xi_{j_2}]\mathbb{E}[\xi_{j_1}] + 2\,(\mathbb{E}[\xi_{j_1}])^2\mathbb{E}[\xi_{j_2}] \\
&\quad = x_{j_1}x_{j_2}\left[m^{(2)} + m^{(3)}x_{j_1}\right] - x_{j_1}\left[m + m^{(2)}x_{j_1}\right]mx_{j_2} - 2m^{(2)}x_{j_1}x_{j_2}mx_{j_1} + 2m^2x_{j_1}^2mx_{j_2} \\
&\quad = \boxed{mx_{j_1}x_{j_2}(2x_{j_1} - 1)} \qquad (80)
\end{aligned}$$

$$\begin{aligned}
&\mathbb{E}[(\xi_{j_1} - \mathbb{E}[\xi_{j_1}])(\xi_{j_2} - \mathbb{E}[\xi_{j_2}])(\xi_{j_3} - \mathbb{E}[\xi_{j_3}])] \\
&\quad = \mathbb{E}[\xi_{j_1}\xi_{j_2}\xi_{j_3}] - \mathbb{E}[\xi_{j_1}\xi_{j_2}]\mathbb{E}[\xi_{j_3}] - \mathbb{E}[\xi_{j_1}\xi_{j_3}]\mathbb{E}[\xi_{j_2}] - \mathbb{E}[\xi_{j_2}\xi_{j_3}]\mathbb{E}[\xi_{j_1}] + 2\,\mathbb{E}[\xi_{j_1}]\mathbb{E}[\xi_{j_2}]\mathbb{E}[\xi_{j_3}] \\
&\quad = m^{(3)}x_{j_1}x_{j_2}x_{j_3} - m^{(2)}x_{j_1}x_{j_2}mx_{j_3} - m^{(2)}x_{j_1}x_{j_3}mx_{j_2} - m^{(2)}x_{j_2}x_{j_3}mx_{j_1} + 2m^3x_{j_1}x_{j_2}x_{j_3} \\
&\quad = \boxed{2mx_{j_1}x_{j_2}x_{j_3}}. \qquad (81)
\end{aligned}$$

<u>Order 4:</u> For any distinct $j_1, j_2, j_3, j_4 \in \{1, 2, \dots, d\}$,

$$\begin{aligned}
&\mathbb{E}[(\xi_{j_1} - \mathbb{E}[\xi_{j_1}])^4] \\
&\quad = \mathbb{E}[\xi_{j_1}^4] - 4\,\mathbb{E}[\xi_{j_1}^3]\mathbb{E}[\xi_{j_1}] + 6\,\mathbb{E}[\xi_{j_1}^2](\mathbb{E}[\xi_{j_1}])^2 - 3\,(\mathbb{E}[\xi_{j_1}])^4 \\
&\quad = x_{j_1}\left[m + 7m^{(2)}x_{j_1} + 6m^{(3)}x_{j_1}^2 + m^{(4)}x_{j_1}^3\right] - 4x_{j_1}\left[m + 3m^{(2)}x_{j_1} + m^{(3)}x_{j_1}^2\right]mx_{j_1} \\
&\qquad + 6x_{j_1}\left[m + m^{(2)}x_{j_1}\right](mx_{j_1})^2 - 3m^4x_{j_1}^4 \\
&\quad = \boxed{3m^2x_{j_1}^2(x_{j_1} - 1)^2 + mx_{j_1}(1 - x_{j_1})(6x_{j_1}^2 - 6x_{j_1} + 1)} \qquad (82)
\end{aligned}$$

$$\begin{aligned}
&\mathbb{E}[(\xi_{j_1} - \mathbb{E}[\xi_{j_1}])^3(\xi_{j_2} - \mathbb{E}[\xi_{j_2}])] \\
&\quad = \mathbb{E}[\xi_{j_1}^3\xi_{j_2}] - \mathbb{E}[\xi_{j_1}^3]\mathbb{E}[\xi_{j_2}] - 3\,\mathbb{E}[\xi_{j_1}^2\xi_{j_2}]\mathbb{E}[\xi_{j_1}] + 3\,\mathbb{E}[\xi_{j_1}^2]\mathbb{E}[\xi_{j_1}]\mathbb{E}[\xi_{j_2}] \\
&\qquad + 3\,\mathbb{E}[\xi_{j_1}\xi_{j_2}](\mathbb{E}[\xi_{j_1}])^2 - 3\,(\mathbb{E}[\xi_{j_1}])^3\mathbb{E}[\xi_{j_2}] \\
&\quad = x_{j_1}x_{j_2}\left[m^{(2)} + 3m^{(3)}x_{j_1} + m^{(4)}x_{j_1}^2\right] - x_{j_1}\left[m + 3m^{(2)}x_{j_1} + m^{(3)}x_{j_1}^2\right]mx_{j_2} \\
&\qquad - 3x_{j_1}x_{j_2}\left[m^{(2)} + m^{(3)}x_{j_1}\right]mx_{j_1} + 3x_{j_1}\left[m + m^{(2)}x_{j_1}\right]mx_{j_1}mx_{j_2} + 3m^{(2)}x_{j_1}x_{j_2}m^2x_{j_1}^2 - 3m^3x_{j_1}^3mx_{j_2} \\
&\quad = \boxed{mx_{j_1}x_{j_2}(3(m-2)x_{j_1}(x_{j_1} - 1) - 1)} \qquad (83)
\end{aligned}$$

$$\begin{aligned}
&\mathbb{E}[(\xi_{j_1} - \mathbb{E}[\xi_{j_1}])^2(\xi_{j_2} - \mathbb{E}[\xi_{j_2}])^2] \\
&\quad = \mathbb{E}[\xi_{j_1}^2\xi_{j_2}^2] - 2\,\mathbb{E}[\xi_{j_1}^2\xi_{j_2}]\mathbb{E}[\xi_{j_2}] - 2\,\mathbb{E}[\xi_{j_1}\xi_{j_2}^2]\mathbb{E}[\xi_{j_1}] + \mathbb{E}[\xi_{j_1}^2](\mathbb{E}[\xi_{j_2}])^2 + \mathbb{E}[\xi_{j_2}^2](\mathbb{E}[\xi_{j_1}])^2 \\
&\qquad + 4\,\mathbb{E}[\xi_{j_1}\xi_{j_2}]\mathbb{E}[\xi_{j_1}]\mathbb{E}[\xi_{j_2}] - 3\,(\mathbb{E}[\xi_{j_1}])^2(\mathbb{E}[\xi_{j_2}])^2 \\
&\quad = x_{j_1}x_{j_2}\left[m^{(2)} + m^{(3)}(x_{j_1} + x_{j_2}) + m^{(4)}x_{j_1}x_{j_2}\right] - 2x_{j_1}x_{j_2}\left[m^{(2)} + m^{(3)}x_{j_1}\right]mx_{j_2} \\
&\qquad - 2x_{j_1}x_{j_2}\left[m^{(2)} + m^{(3)}x_{j_2}\right]mx_{j_1} + x_{j_1}\left[m + m^{(2)}x_{j_1}\right]m^2x_{j_2}^2 + x_{j_2}\left[m + m^{(2)}x_{j_2}\right]m^2x_{j_1}^2 \\
&\qquad + 4m^{(2)}x_{j_1}x_{j_2}mx_{j_1}mx_{j_2} - 3\,m^2x_{j_1}^2m^2x_{j_2}^2 \\
&\quad = \boxed{m(m-2)x_{j_1}x_{j_2}(3x_{j_1}x_{j_2} - (x_{j_1} + x_{j_2}) + 1) + mx_{j_1}x_{j_2}} \qquad (84)
\end{aligned}$$

$$\begin{aligned}
&\mathbb{E}[(\xi_{j_1} - \mathbb{E}[\xi_{j_1}])^2(\xi_{j_2} - \mathbb{E}[\xi_{j_2}])(\xi_{j_3} - \mathbb{E}[\xi_{j_3}])] \\
&\quad = \mathbb{E}[\xi_{j_1}^2\xi_{j_2}\xi_{j_3}] - \mathbb{E}[\xi_{j_1}^2\xi_{j_2}]\mathbb{E}[\xi_{j_3}] - \mathbb{E}[\xi_{j_1}^2\xi_{j_3}]\mathbb{E}[\xi_{j_2}] - 2\,\mathbb{E}[\xi_{j_1}\xi_{j_2}\xi_{j_3}]\mathbb{E}[\xi_{j_1}] + \mathbb{E}[\xi_{j_1}^2]\mathbb{E}[\xi_{j_2}]\mathbb{E}[\xi_{j_3}]
\end{aligned}$$

$$
\begin{aligned}
&\quad + 2\,\mathbb{E}[\xi_{j_1}\xi_{j_2}]\mathbb{E}[\xi_{j_1}]\mathbb{E}[\xi_{j_3}] + 2\,\mathbb{E}[\xi_{j_1}\xi_{j_3}]\mathbb{E}[\xi_{j_1}]\mathbb{E}[\xi_{j_2}] + \mathbb{E}[\xi_{j_2}\xi_{j_3}](\mathbb{E}[\xi_{j_1}])^2 - 3\,(\mathbb{E}[\xi_{j_1}])^2\mathbb{E}[\xi_{j_2}]\mathbb{E}[\xi_{j_3}] \\
&= x_{j_1}x_{j_2}x_{j_3}\,[m^{(3)} + m^{(4)}x_{j_1}] - x_{j_1}x_{j_2}\,[m^{(2)} + m^{(3)}x_{j_1}]mx_{j_3} - x_{j_1}x_{j_3}\,[m^{(2)} + m^{(3)}x_{j_1}]mx_{j_2} \\
&\quad - 2m^{(3)}x_{j_1}x_{j_2}x_{j_3}mx_{j_1} + x_{j_1}\,[m + m^{(2)}x_{j_1}]\,mx_{j_2}mx_{j_3} + 2m^{(2)}x_{j_1}x_{j_2}mx_{j_1}mx_{j_3} \\
&\quad + 2m^{(2)}x_{j_1}x_{j_3}mx_{j_1}mx_{j_2} + m^{(2)}x_{j_2}x_{j_3}m^2x_{j_1}^2 - 3m^2x_{j_1}^2mx_{j_2}mx_{j_3} \\
&= \boxed{m(m-2)x_{j_1}x_{j_2}x_{j_3}(3x_{j_1}-1)}
\end{aligned} \tag{85}
$$

$$
\begin{aligned}
&\mathbb{E}[(\xi_{j_1} - \mathbb{E}[\xi_{j_1}])(\xi_{j_2} - \mathbb{E}[\xi_{j_2}])(\xi_{j_3} - \mathbb{E}[\xi_{j_3}])(\xi_{j_4} - \mathbb{E}[\xi_{j_4}])] \\
&\quad = \mathbb{E}[\xi_{j_1}\xi_{j_2}\xi_{j_3}\xi_{j_4}] - \mathbb{E}[\xi_{j_1}\xi_{j_2}\xi_{j_3}]\mathbb{E}[\xi_{j_4}] - \mathbb{E}[\xi_{j_1}\xi_{j_2}\xi_{j_4}]\mathbb{E}[\xi_{j_3}] - \mathbb{E}[\xi_{j_1}\xi_{j_3}\xi_{j_4}]\mathbb{E}[\xi_{j_2}] - \mathbb{E}[\xi_{j_2}\xi_{j_3}\xi_{j_4}]\mathbb{E}[\xi_{j_1}] \\
&\quad\quad + \mathbb{E}[\xi_{j_1}\xi_{j_2}]\mathbb{E}[\xi_{j_3}]\mathbb{E}[\xi_{j_4}] + \mathbb{E}[\xi_{j_1}\xi_{j_3}]\mathbb{E}[\xi_{j_2}]\mathbb{E}[\xi_{j_4}] + \mathbb{E}[\xi_{j_1}\xi_{j_4}]\mathbb{E}[\xi_{j_2}]\mathbb{E}[\xi_{j_3}] \\
&\quad\quad + \mathbb{E}[\xi_{j_2}\xi_{j_3}]\mathbb{E}[\xi_{j_1}]\mathbb{E}[\xi_{j_4}] + \mathbb{E}[\xi_{j_2}\xi_{j_4}]\mathbb{E}[\xi_{j_1}]\mathbb{E}[\xi_{j_3}] + \mathbb{E}[\xi_{j_3}\xi_{j_4}]\mathbb{E}[\xi_{j_1}]\mathbb{E}[\xi_{j_2}] - 3\,\mathbb{E}[\xi_{j_1}]\mathbb{E}[\xi_{j_2}]\mathbb{E}[\xi_{j_3}]\mathbb{E}[\xi_{j_4}] \\
&\quad = m^{(4)}x_{j_1}x_{j_2}x_{j_3}x_{j_4} - m^{(3)}x_{j_1}x_{j_2}x_{j_3}mx_{j_4} - m^{(3)}x_{j_1}x_{j_2}x_{j_4}mx_{j_3} - m^{(3)}x_{j_1}x_{j_3}x_{j_4}mx_{j_2} - m^{(3)}x_{j_2}x_{j_3}x_{j_4}mx_{j_1} \\
&\quad\quad + m^{(2)}x_{j_1}x_{j_2}mx_{j_3}mx_{j_4} + m^{(2)}x_{j_1}x_{j_3}mx_{j_2}mx_{j_4} + m^{(2)}x_{j_1}x_{j_4}mx_{j_2}mx_{j_3} \\
&\quad\quad + m^{(2)}x_{j_2}x_{j_3}mx_{j_1}mx_{j_4} + m^{(2)}x_{j_2}x_{j_4}mx_{j_1}mx_{j_3} + m^{(2)}x_{j_3}x_{j_4}mx_{j_1}mx_{j_2} - 3\,m^4x_{j_1}x_{j_2}x_{j_3}x_{j_4} \\
&\quad = \boxed{3m(m-2)x_{j_1}x_{j_2}x_{j_3}x_{j_4}}\,.
\end{aligned} \tag{86}
$$

6. Conclusions

In this short paper, we found general formulas for the central and non-central moments of the multinomial distribution, as well as explicit formulas for all the non-central moments up to order 8 and all the central moments up to order 4. Our work expands on the results in [2], where the central moments were calculated up to order 4. It also complements the general formula for the (joint) factorial moments from [1] and the explicit formulas for some of the lower-order (mixed) cumulants that were presented in [3].

Author Contributions: All contributions were made by the sole author of the article, F.O. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by a postdoctoral fellowship from the NSERC (PDF) and a postdoctoral fellowship supplement from the FRQNT (B3X).

Acknowledgments: We thank the three referees for their comments.

Conflicts of Interest: The author declares no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

i.i.d.	independent and identically distributed

References

1. Mosimann, J.E. On the compound multinomial distribution, the multivariate β-distribution, and correlations among proportions. *Biometrika* **1962**, *49*, 65–82.
2. Newcomer, J.T. Estimation Procedures for Multinomial Models with Overdispersion. Ph.D. Thesis, University of Maryland, College Park, MD, USA, 2008.
3. Wishart, J. Cumulants of multivariate multinomial distributions. *Biometrika* **1949**, *36*, 47–58.
4. Johnson, N.L.; Kotz, S.; Balakrishnan, N. *Discrete Multivariate Distributions*; Wiley Series in Probability and Statistics: Applied Probability and Statistics; John Wiley & Sons, Inc.: New York, NY, USA, 1997; pp. xxii+299.
5. Vitale, R.A. Bernstein polynomial approach to density function estimation. In *Statistical Inference and Related Topics*; Academic Press: New York, NY, USA, 1975; pp. 87–99.
6. Gawronski, W.; Stadtmüller, U. Smoothing histograms by means of lattice and continuous distributions. *Metrika* **1981**, *28*, 155–164.
7. Stadtmüller, U. Asymptotic distributions of smoothed histograms. *Metrika* **1983**, *30*, 145–158.
8. Gawronski, W. Strong laws for density estimators of Bernstein type. *Period. Math. Hungar* **1985**, *16*, 23–43.
9. Stadtmüller, U. Asymptotic properties of nonparametric curve estimates. *Period. Math. Hungar.* **1986**, *17*, 83–108.
10. Tenbusch, A. Nonparametric curve estimation with Bernstein estimates. *Metrika* **1997**, *45*, 1–30.

11. Petrone, S. Bayesian density estimation using Bernstein polynomials. *Can. J. Stat.* **1999**, *27*, 105–126.
12. Petrone, S. Random Bernstein polynomials. *Scand. J. Stat.* **1999**, *26*, 373–393.
13. Ghosal, S. Convergence rates for density estimation with Bernstein polynomials. *Ann. Stat.* **2001**, *29*, 1264–1280.
14. Petrone, S.; Wasserman, L. Consistency of Bernstein polynomial posteriors. *J. Roy. Stat. Soc. Ser. B* **2002**, *64*, 79–100.
15. Babu, G.J.; Canty, A.J.; Chaubey, Y.P. Application of Bernstein polynomials for smooth estimation of a distribution and density function. *J. Stat. Plan. Inference* **2002**, *105*, 377–392.
16. Kakizawa, Y. Bernstein polynomial probability density estimation. *J. Nonparametric Stat.* **2004**, *16*, 709–729.
17. Bouezmarni, T.; Rolin, J.M. Bernstein estimator for unbounded density function. *J. Nonparametric Stat.* **2007**, *19*, 145–161.
18. Bouezmarni, T.; M., M.; Rolin, J.M. L_1-rate of convergence of smoothed histogram. *Stat. Probab. Lett.* **2007**, *77*, 1497–1504.
19. Leblanc, A. Chung-Smirnov property for Bernstein estimators of distribution functions. *J. Nonparametric Stat.* **2009**, *21*, 133–142.
20. Leblanc, A. A bias-reduced approach to density estimation using Bernstein polynomials. *J. Nonparametric Stat.* **2010**, *22*, 459–475.
21. Curtis, S.M.; Ghosh, S.K. A variable selection approach to monotonic regression with Bernstein polynomials. *J. Appl. Stat.* **2011**, *38*, 961–976.
22. Leblanc, A. On estimating distribution functions using Bernstein polynomials. *Ann. Inst. Stat. Math.* **2012**, *64*, 919–943.
23. Leblanc, A. On the boundary properties of Bernstein polynomial estimators of density and distribution functions. *J. Stat. Plan. Inference* **2012**, *142*, 2762–2778.
24. Igarashi, G.; Kakizawa, Y. On improving convergence rate of Bernstein polynomial density estimator. *J. Nonparametric Stat.* **2014**, *26*, 61–84.
25. Turnbull, B.C.; Ghosh, S.K. Unimodal density estimation using Bernstein polynomials. *Comput. Stat. Data Anal.* **2014**, *72*, 13–29.
26. Lu, L. On the uniform consistency of the Bernstein density estimator. *Stat. Probab. Lett.* **2015**, *107*, 52–61.
27. Guan, Z. Efficient and robust density estimation using Bernstein type polynomials. *J. Nonparametric Stat.* **2016**, *28*, 250–271.
28. Guan, Z. Bernstein polynomial model for grouped continuous data. *J. Nonparametric Stat.* **2017**, *29*, 831–848.
29. Belalia, M.; Bouezmarni, T.; Leblanc, A. Smooth conditional distribution estimators using Bernstein polynomials. *Comput. Stat. Data Anal.* **2017**, *111*, 166–182.
30. Belalia, M.; Bouezmarni, T.; Leblanc, A. Bernstein conditional density estimation with application to conditional distribution and regression functions. *J. Korean Stat. Soc.* **2019**, *48*, 356–383.
31. Tenbusch, A. Two-dimensional Bernstein polynomial density estimators. *Metrika* **1994**, *41*, 233–253.
32. Ouimet, F. Asymptotic properties of Bernstein estimators on the simplex. *arXiv* **2020**, arXiv:2002.07758.
33. Ouimet, F. Asymptotic properties of Bernstein estimators on the simplex. Part 2: the boundary case. *arXiv* **2020**, arXiv:2006.11756.
34. Ouimet, F. A precise local limit theorem for the multinomial distribution and some applications. *arXiv* **2020**, arXiv:2001.08512.
35. Arenbaev, N.K. Asymptotic behavior of the multinomial distribution. *Teor. Veroyatnost. Primenen.* **1976**, *21*, 826–831.
36. Aitchison, J.; Lauder, I.J. Kernel Density Estimation for Compositional Data. *J. R. Stat. Soc. Ser. (Appl. Stat.)* **1985**, *34*, 129–137, doi:10.2307/2347365.
37. Brown, B.M.; Chen, S.X. Beta-Bernstein smoothing for regression curves with compact support. *Scand. J. Stat.* **1999**, *26*, 47–59.
38. Chen, S.X. Beta kernel estimators for density functions. *Comput. Stat. Data Anal.* **1999**, *31*, 131–145.
39. Chen, S.X. Beta kernel smoothers for regression curves. *Stat. Sin.* **2000**, *10*, 73–91.
40. Bouezmarni, T.; Rolin, J.M. Consistency of the beta kernel density function estimator. *Can. J. Stat.* **2003**, *31*, 89–98.
41. Ouimet, F. Density estimation using Dirichlet kernels. *arXiv* **2020**, arXiv:2002.06956.
42. Graham, R.L.; Knuth, D.E.; Patashnik, O. *Concrete Mathematics*, 2nd ed.; Addison-Wesley Publishing Company: Reading, MA, USA, 1994; pp. xiv+657.
43. Newcomer, J.T.; Neerchal, N.K.; Morel, J.G. Computation of Higher Order Moments from Two Multinomial Overdispersion Likelihood Models. 2008; pp. 1–11. Available online: http://www.math.umbc.edu/~kogan/technical_papers/2008/Newcomer_Nagaraj_Morel.pdf (accessed on 10 November 2020).
44. Ouimet, F. Explicit formulas for the joint third and fourth central moments of the multinomial distribution. *arXiv* **2020**, arXiv:2006.09059.